I Claim:

1. A method for determining the sequence of a polynucleotide, comprising

a. providing a nucleic acid fragment having homology to a known reference sequence, and

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the sequence of said fragment by comparing said at least one property to the predicted properties of polypeptides encoded in said known reference sequence.

2. The method of claim 1 wherein said nucleic acid fragment contains a difference with respect to the reference sequence wherein said difference is selected from the group consisting of single nucleotide polymorphism, single nucleotide substitution, single nucleotide deletion, single nucleotide insertion, multiple nucleotide substitution, multiple nucleotide deletion, multiple nucleotide insertion, DNA duplication, DNA inversion, DNA translocation, and DNA deletion/substitution.

3. The method of claim 1 wherein said nucleic acid fragment comprises an exon.

4. The method of claim 1 wherein said nucleic acid fragment comprises a cDNA.

5. The method of claim 1 wherein said at least one polypeptide comprises a fragment homologous to said reference sequence and at least one predetermined heterologous epitope tag.

6. The method of claim 1 wherein said at least one polypeptide is expressed in a living cell.

7. The method of claim 1 wherein said at least one polypeptide is expressed in a cell free system.

8. The method of claim 7 wherein said cell free system is selected from the group consisting of E. coli extract, rabbit reticulocyte extract, and wheat germ extract.

9. The method of claim 1 further comprising purifying said peptide in conjunction with assessing the physical property.

10. The method of claim 9 wherein said purification comprises a method selected from the group consisting of gel electrophoresis, capillary electrophoresis, liquid chromatography (LC), capillary liquid chromatography, high performance liquid chromatography (HPLC), differential centrifugation, filtration, gel filtration, membrane chromatography, affinity purification, biomolecular interaction analysis (BIA), ligand affinity purification, glutathione-S-transferase affinity chromatography, cellulose binding protein affinity chromatography, maltose binding protein affinity chromatography, avidin/streptavidin affinity chromatography, S-tag affinity chromatography, thioredoxin affinity chromatography, metal-chelate affinity chromatography, immobilized metal affinity chromatography, epitope-tag affinity chromatography, immunoaffinity chromatography, immunoaffinity capture, capture using bioreactive mass spectrometer probes, mass spectrometric immunoassay, and immunoprecipitation.

11. The method of claim 1 wherein the physical property that is determined is mass.

12. The method of claim 11 wherein said mass is determined by a method selected from the group consisting of mass spectrometry, MALDI-TOF mass spectrometry, electrospray ionization mass spectrometry (ESI) ) tandem mass spectrometry (MS/MS), quadripole time of flight spectrometry (Q-TOF), Fourier transform ion cyclotron resonance (FTICR) mass spectrometry, gel electrophoresis, capillary electrophoresis, and high performance liquid chromatography (HPLC).

13. The method of claim 1 wherein the physical property that is assessed is partial or complete amino acid composition.

14. The method of claim 1 wherein the physical property that is assessed is partial or complete amino acid sequence.

15. A method for genetic analysis, comprising

a. providing a nucleic acid fragment having homology to a known reference sequence, and

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the coding capacity of said fragment by comparing said at least one physical property to the predicted properties of polypeptides encoded in a known reference sequence.

16. The method of claim 15 wherein said nucleic acid fragment contains a difference with respect to the reference sequence selected from the group consisting of single nucleotide polymorphism, single nucleotide substitution, single nucleotide deletion, single nucleotide insertion, multiple nucleotide substitution, multiple nucleotide deletion, multiple nucleotide insertion, DNA duplication, DNA inversion, DNA translocation, and DNA deletion/substitution.

17. The method of claim 15 wherein said nucleic acid fragment comprises an exon.

18. The method of claim 15 wherein said nucleic acid fragment comprises a cDNA.

19. The method of claim 15 wherein said at least one polypeptide contains at least one epitope tag.

20. The method of claim 15 wherein said at least one polypeptide is expressed in a living cell.

21. The method of claim 15 wherein said at least one polypeptide is expressed in a cell free system.

22. The method of claim 21 wherein said cell free system is selected from the group consisting of E. coli extract, rabbit reticulocyte extract, and wheat germ extract.

23. The method of claim 15 further comprising purification of said peptide in conjunction with assessing the physical property.

24. The method of claim 23 wherein said purification comprises a method selected from the group consisting of gel electrophoresis, capillary electrophoresis, liquid chromatography (LC), capillary liquid chromatography, high performance liquid chromatography (HPLC), differential centrifugation, filtration, gel filtration, membrane chromatography, affinity purification, biomolecular interaction analysis (BIA), ligand affinity purification, glutathione-S-transferase affinity chromatography, cellulose binding protein affinity chromatography, maltose binding protein affinity chromatography, avidin/streptavidin affinity chromatography, S-tag affinity chromatography, thioredoxin affinity chromatography, metal-chelate affinity chromatography, immobilized metal affinity chromatography, epitope-tag affinity chromatography, immunoaffinity chromatography, immunoaffinity capture, capture using bioreactive mass spectrometer probes, mass spectrometric immunoassay, and immunoprecipitation.

25. The method of claim 15 wherein the physical property that is determined is mass.

26. The method of claim 25 wherein said mass is determined by a method selected from the group consisting of mass spectrometry, MALDI-TOF mass spectrometry, electrospray ionization mass spectrometry (ESI) tandem mass

spectrometry (MS/MS), quadripole time of flight spectrometry (Q-TOF), Fourier transform ion cyclotron resonance (FTICR) mass spectrometry, gel electrophoresis, capillary electrophoresis, and high performance liquid chromatography (HPLC).

27. The method of claim 15 wherein the physical property that is assessed is partial or complete amino acid composition.

28. The method of claim 15 wherein the physical property that is assessed is partial or complete amino acid sequence.

29. A method for assessing a disease, condition, genotype, or phenotype, comprising

a. providing a nucleic acid fragment from a biological sample, and

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the sequence of said fragment by comparing said at least one property to the predicted properties of polypeptides encoded in a known reference sequence, and

d. correlating said determined sequence with said disease, condition, genotype or phenotype.

30. The method of claim 29 wherein the original source of said biological sample is obtained from a virus, organelle, cell, tissue, body part, exudate, excretion, elimination, secretion, blood, sweat, urine, tears, semen, saliva, feces, skin, hair or milk of a healthy, diseased or deceased microorganism, protist, alga, fungus, animal or plant.

31. A diagnostic or prognostic test for a disease, condition, genotype, or phenotype, comprising

a. providing a nucleic acid fragment from a biological sample,

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the sequence of said fragment by comparing said at least one property to the predicted properties of polypeptides encoded in said known reference sequence.

32. The method of claim 31 wherein the original source of said biological sample is a virus, organelle, cell, tissue, body part, exudate, excretion, elimination, secretion, blood, sweat, urine, tears, semen, saliva, feces, skin, hair or milk of a healthy, diseased or deceased microorganism, protist, alga, fungus, animal or plant.

33. The diagnostic or prognostic test of claim 31 wherein said test detects heterozygote status.

34. The diagnostic or prognostic test of claim 31 wherein said phenotype is response to a drug or therapeutic treatment.

35. The diagnostic or prognostic test of claim 31 wherein said disease is a genetic disease.

36. The diagnostic or prognostic test of claim 31 wherein the genetic disease is selected from the group consisting of Alzheimer's disease, Ataxia talangietasia (ATM), Familial ademato us polyposis (APC), Hereditary breast/ovarian cancer (BRCA1, BRCA2), Hereditary melanoma (CDK2, CDKN2), Hereditary non-polypsosis colon cancer (hMSH2, hMLH1, hPMS1, hPMS2), Hereditary retinoblastoma (RB1), Hereditary Wilm's Tumor (WT1), Li-Fraumeni syndrome (p53), Multiple endocrine neoplasia (MEN1, MEN2), Von Hippel-Lindau syndrome (VHL), Congenital adrenal hyperplasia, Androgen receptor deficiency, Tetrahydrobiopterin deficiency, X-Linked agammaglobulinemia, Cystic Fibrosis (CFTR), Diabetes, Muscular Dystrophy (DMD, BMD), Factor X deficiency, Mitochondrial gene deficiency, Factor VII deficiency, Glucose-6-Phosphate deficiency, Pompe Disease, Hemophilia A, Hexosaminidase A deficiency, Human Type I and Type III Collagen deficiency X-linked SCID, Retinitis pigmentosa (RP)

LIACAM deficiency, MCAD deficiency, LDL Receptor deficiency, Ornithine Transcarbamylase deficiency, PAX6 Mutation Phenylketonuria, RB1 Gene Mutation, Tuberous Sclerosis, von Willebrand Factor Disease, and Werner Syndrome.

37. The diagnostic or prognostic test of claim 31 wherein said disease is cancer.

38. The diagnostic or prognostic test of claim 31 wherein said disease is an infectious disease.

39. A method for assessing a disease, condition, genotype, or phenotype, comprising

a. providing a nucleic acid fragment from a biological sample,

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the coding capacity of said fragment by comparing said at least one property to the predicted properties of polypeptides encoded in a known reference sequence.

d. correlating said determined sequence with said disease, condition, genotype or phenotype.

40. The method of claim 39 wherein the original source of said biological sample is a virus, organelle, cell, tissue, body part, exudate, excretion, elimination, secretion, blood, sweat, urine, tears, semen, saliva, feces, skin, hair or milk of a healthy, diseased or deceased microorganism, protist, alga, fungus, animal or plant.

41. A diagnostic or prognostic test for a disease, condition, genotype, or phenotype, comprising

a. providing a nucleic acid fragment from a biological sample,

b. expressing at least one polypeptide from said fragment, and

c. assessing at least one physical property of said at least one polypeptide to determine the coding capacity of said fragment by comparing said at

least one property to the predicted properties of polypeptides encoded in a known reference sequence.

42. The test of claim 41 wherein the original source of said biological sample is a virus, organelle, cell, tissue, body part, exudate, excretion, elimination, secretion, blood, sweat, urine, tears, semen, saliva, feces, skin, hair or milk of a healthy, diseased or deceased microorganism, protist, alga, fungus, animal or plant.

43. The diagnostic or prognostic test of claim 41 wherein said test detects heterozygote status.

44. The diagnostic or prognostic test of claim 41 wherein said phenotype is response to a therapeutic drug or treatment.

45. The diagnostic or prognostic test of claim 41 wherein said disease is a genetic disease.

46. The diagnostic or prognostic test of claim 41 wherein the genetic disease is selected from the group consisting of Alzheimer's disease, Ataxia talangietasia (ATM), Familial ademato us polyposis (APC), Hereditary breast/ovarian cancer (BRCA1, BRCA2), Hereditary melanoma (CDK2, CDKN2), Hereditary non-polypsosis colon cancer (hMSH2, hMLH1, hPMS1, hPMS2), Hereditary retinoblastoma (RB1), Hereditary Wilm's Tumor (WT1), Li-Fraumeni syndrome (p53), Multiple endocrine neoplasia (MEN1, MEN2), Von Hippel-Lindau syndrome (VHL), Congenital adrenal hyperplasia, Androgen receptor deficiency, Tetrahydrobiopterin deficiency, X-Linked agammaglobulinemia, Cystic Fibrosis (CFTR), Diabetes, Muscular Dystrophy (DMD, BMD), Factor X deficiency, Mitochondrial gene deficiency, Factor VII deficiency, Glucose-6-Phosphate deficiency, Pompe Disease, Hemophilia A, Hexosaminidase A deficiency, Human Type I and Type III Collagen deficiency X-linked SCID, Retinitis pigmentosa (RP) LIACAM deficiency, MCAD deficiency, LDL Receptor deficiency, Ornithine

Transcarbamylase deficiency, PAX6 Mutation Phenylketonuria, RB1 Gene Mutation, Tuberous Sclerosis, von Willebrand Factor Disease, and Werner Syndrome.

47. The diagnostic or prognostic test of claim 41 wherein said disease is cancer.

48. The diagnostic or prognostic test of claim 41 wherein said disease is an infectious disease.

49. Said at least one polypeptide of claim 1.

50. Said at least one polypeptide of claim 15.

51. Said at least one polypeptide of claim 29.

52. Said at least one polypeptide of claim 31.

53. Said at least one polypeptide of claim 39.

54. Said at least one polypeptide of claim 41.

55. A data structure useful for detecting and analyzing DNA mutations and polymorphisms, comprising:

a. data representing the following stored in a physical medium in computer readable form:

i. a plurality of DNA sequence fragments contained within a reference DNA sequence, and

ii. the sequences of the polypeptides encoded in said DNA sequence fragments, and

iii. the predicted sequences of a plurality of polypeptides encoded in a set of transformed DNA sequence fragments, each member of said set comprised of a DNA sequence related to said DNA sequence fragment by a specific change selected from the group consisting of single nucleotide polymorphism, single

nucleotide substitution, single nucleotide deletion, single nucleotide insertion, multiple nucleotide substitution, multiple nucleotide deletion, multiple nucleotide insertion, DNA duplication, DNA inversion, DNA translocation, and DNA deletion/substitution;

b. means for comparing the predicted sequences of said plurality of polypeptides with a test sequence to determine identity of the test sequence with a predicted sequence.

56. A computer data structure, comprising:

a data storage medium;

a first data set in computer readable form on said data storage medium representing a plurality of polypeptide fragments of a polypeptide encoded by a reference polynucleotide sequence;

a second data set in computer readable form on said data storage medium representing a physical property of each of said polypeptide fragments; and

means for correlating an empirically derived physical property of a test polypeptide with the second data set to determine the identity of the test polypeptide.

57. The data structure of claim 56 further comprising a third data set in computer readable form on said data storage medium representing polynucleotide fragments encoding said polypeptide fragments; and means for correlating the identity of the test polypeptide with a polynucleotide fragment represented in said third data set.

58. The data structure of claim 56 wherein said physical property is mass.

59. The data structure of claim 56 wherein said reference polynucleotide has a reading frame, and wherein said first data set represents polypeptide fragments encoded in frame and polypeptide fragments encoded out of frame with respect to said reference polynucleotide.

60. A computer implemented method for ascertaining the identity of a nucleic acid fragment encoding a polypeptide, wherein the nucleic acid fragment is a fragment of a known reference sequence, comprising the steps of:

measuring a physical property of said polypeptide;

comparing, in a computer, the measured physical property with a data set representing the predicted corresponding physical properties of possible polypeptides that are encoded by fragments of said reference sequence within a predetermined size range;

identifying a match between said measured physical property and a predicted physical property in the data set; and

displaying or recording the results of the identifying step.

61. The method of claim 60 wherein said reference polynucleotide has a frame, and said data set includes physical properties of polypeptides encoded by out-of-frame fragments of said reference polynucleotide.

62. The method of claim 60 wherein said reference polynucleotide has six possible frames, and said data set includes physical properties of polypeptides encoded by fragments having at least one of said possible frames.

63. A relational data set useful for detecting and analyzing DNA mutations and polymorphisms comprising,

a. a plurality of DNA sequence fragments contained within a reference DNA sequence,

b. the sequences of the polypeptides encoded in said DNA sequence fragments, and

c. the predicted sequences of a plurality of polypeptides encoded in a set of transformed DNA sequence fragments, each member of said set comprised of a DNA sequence related to said DNA sequence fragment by a specific change selected from the group consisting of single nucleotide polymorphism, single nucleotide substitution, single nucleotide deletion, single nucleotide insertion, multiple nucleotide substitution,

multiple nucleotide deletion, multiple nucleotide insertion, DNA duplication, DNA inversion, DNA translocation, and DNA deletion/substitution.

64. A computer program comprising a search of the data set of claim 63.

65. A method for genetic analysis, comprising:

providing two or more nucleic acid samples derived from two or more biological samples, said biological samples being heterogeneous;

expressing polypeptides from each of said nucleic acid samples;

subjecting said polypeptides, in combination, to physical property assessment; and

comparing the results of said physical property assessment to predicted properties of polypeptides encoded in at least one known reference sequence.

66. The method according to claim 65 wherein said nucleic acid fragments are derived by PCR.

67. The method according to claim 66 wherein a different PCR primer is selected for each biological specimen.

68. The method according to claim 67 wherein each different PCR primer is identical in its 3' portion and differs at its 5' portion.

69. The method The method according to claim 67 wherein each different PCR primer is identical in its 5' portion and differs at its 3' portion.

70. The method according to claim 66 wherein said PCR amplicons are physically distinguishable, as are the peptides that they encode.

71. The method according to claim 65 wherein the heterogeneity of said biological sample is attributable to said samples having derived from different individuals.

72. The method according to claim 65 wherein the heterogeneity of said biological samples is attributable to said samples having derived from heterogeneous tissue from a single individual.

73. The method according to claim 65 wherein the heterogeneity of said biological samples is attributable to said samples having derived from a heterozygous cell or individual.

74. The method according to claim 1, wherein the step of expressing at least one polypeptide from said fragment is performed in a nonsense-suppressing environment.

75. The method according to claim 1, wherein the step of expressing at least one polypeptide from said fragment is performed in a missense-suppressing environment.

76. The method according to claim 15, wherein the step of expressing at least one polypeptide from said fragment is performed in a nonsense-suppressing environment.

77. The method according to claim 15, wherein the step of expressing at least one polypeptide from said fragment is performed in a missense-suppressing environment.

78. The method of claim 9 wherein said peptide is purified by sequential affinity capture by means of more than one distinct affinity element.

79. The method of claim 78 wherein one affinity element resides in the N-terminal portion of said peptide and a second affinity element resides in the C-terminal portion of said peptide.

80. Providing a nucleic acid molecule;

expressing polypeptides from two or more reading frames of said nucleic acid sample; and

determining the masses of said polypeptides to create a peptide mass signature characteristic of said nucleic acid molecule.